



**QUEEN'S
UNIVERSITY
BELFAST**

Off-Centre Localisation Performance of Ambisonics and HOA For Large and Small Loudspeaker Array Radii

Stitt, P., Bertet, S., & Walstijn, M. V. (2014). Off-Centre Localisation Performance of Ambisonics and HOA For Large and Small Loudspeaker Array Radii. *Acta Acustica united with Acustica*, 100(5), 937-944.
<https://doi.org/10.3813/AAA.918773>

Published in:

Acta Acustica united with Acustica

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

The archived file is not the final published version of the article " Stitt, Peter; Bertet, Stéphanie; van Walstijn, Maarten, Off-Centre Localisation Performance of Ambisonics and HOA For Large and Small Loudspeaker Array Radii in: "Acta Acustica united with Acustica", volume 100, number 5, 2014, pp. 937-944. © 2014 S. Hirzel Verlag/European Acoustics Association. The definitive publisher-authenticated version is available online at <http://www.ingentaconnect.com/content/dav/aaua>, <http://dx.doi.org/10.3813/AAA.918773>.

Readers must contact the publisher for reprint or permission to use the material in any form.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

OFF-CENTRE LOCALISATION PERFORMANCE OF AMBISONICS AND HOA FOR LARGE AND SMALL LOUDSPEAKER ARRAY RADII

Peter Stitt, Stéphanie Bertet and Maarten van Walstijn,

Sonic Arts Research Centre
Queen's University, Belfast
Belfast, Northern Ireland

{pstitt01, s.bertet, m.vanwalstijn} @qub.ac.uk

ABSTRACT

Ambisonics and Higher Order Ambisonics (HOA) are scalable spatial audio techniques that attempt to present a sound scene to listeners over as large an area as possible. A localisation experiment was carried out to investigate the performance of a first and third order system at three listening positions - one in the centre and two off-centre - using a 5 m radius loudspeaker array. The results are briefly presented and compared to those of an earlier experiment on a 2.2 m loudspeaker array. In both experiments the off-centre listeners were placed such that the ratio of distance from the centre to the array radius was constant in both experiments. The test used a reverse target-pointer adjustment method to determine the error, both signed and absolute, for each combination of listening position and system. The results for both arrays were found to be very similar, suggesting that the relative amplitude of the loudspeakers, which were the same in both cases, was more dominant for localisation than the arrival time differences, which differed between array sizes.

1. INTRODUCTION

Michael Gerzon [1] first proposed the theory behind Ambisonics in the 1970s as an alternative to the then prevalent quadraphonic systems. It is a multichannel reproduction technique that attempts to recreate a physical sound field over as large a listening area as possible. It is scalable and higher orders allow a larger listening area, but an increase in order also requires more loudspeakers to recreate the sound field [2].

In its most basic form, Ambisonics is used to reconstruct a plane wave by decomposing the sound field using spherical harmonic functions. This process is known as the encoding stage. In theory an infinite number of spherical harmonics must be used to recreate the sound field but in practice the series must be limited to a finite order N . An ambisonic reconstruction of order $N > 1$ is referred to as Higher Order Ambisonics (HOA) [2]. The size of the well reproduced area is frequency dependent and for first and third order is less than the size of a human head above 700 Hz and 1900 Hz respectively. Therefore, without near-field compensation [3] it is an amplitude panning method and at low order gives similar localisation results as other amplitude panning methods [4].

The encoded channels must be decoded to an appropriate loudspeaker array when the sound scene is to be presented to a listener. The decoding process calculates the appropriate loudspeaker gains needed to recreate the sound field and the minimum number of loudspeakers required depends on the order of the reproduction. There are different decoding weightings, such as basic, max r_E [5]

and in-phase [6], which are appropriate for different situations and different weightings can be used over different frequency ranges.

In-phase decoding was proposed for situations where a large number of the listeners cannot be in the sweet spot, such as the presentation of a sound scene to an audience [6]. This is a compromise between the accuracy for a listener at the centre and for those outside the sweet-spot. However, max r_E decoding [5] focusses the energy of the loudspeaker array toward the region of the recreated source. It has been shown to give better localisation performance (the ability of the system to place a sound image) for an off-centre listener than in-phase [7].

Some previous off-centre tests into localisation have been carried out in rooms with moderately long reverberation times and irregular loudspeaker arrangements [7, 8]. In [7] the localisation performance of three decoding weightings was tested for orders 1, 3, and 5, along with a subjective rating of the localisation accuracy. Increasing the order improved the results for the off-centre position and the max r_E decoder was found to give the best results for the setup. In [8] the localisation performance was tested for a number of spatial audio systems at nine listening positions inside the array. At off-centre positions, localisation was biased toward the nearest loudspeaker with all the tested spatialisation systems, especially first order Ambisonics. Increasing the ambisonic order of the system improved the results. This study also had multiple listeners inside the array who would block direct sound to other listeners. However, in both of these experiments the test environments were somewhat reverberant. It therefore becomes difficult to disentangle the influence of the room, the irregularity of the reproduction array and the off-centre listening positions.

In his recent thesis Frank [4] performed a localisation experiment for a central and off-centre listener under more controlled conditions than in [7, 8]. He tested basic and max r_E decoding for a third order system using 8 loudspeakers, along with VBAP [9] and MDAP [10]. The results for the basic decoder showed that the image could split and be perceived as coming from the nearest loudspeakers or from close to the intended image direction. This splitting was not exhibited for the max r_E system. Wierstorf et al. [11] tested Wave Field Synthesis [12] and near-field compensated HOA [3] at multiple off-centre positions using dynamic binaural synthesis. The listeners had to turn their heads to face the perceived direction of the source. Three different loudspeaker spacings were used and it was found for the lowest ambisonic order system tested that the image was drawn in to the nearest loudspeakers at the off-centre listening positions. This study also exhibited some image splitting at listening positions furthest from the centre.

An experiment presented in an earlier paper by the authors [13] was specifically designed to investigate the localisation per-

formance of ambisonic reproduction of two different orders (first and third) at three listening positions, including the centre as reference and the array had a radius of 2.2 m. It was found that the ambisonic image was moved to the further side of the array to compensate for the proximity to the nearest loudspeakers.

It has been noted that the time differences between loudspeaker signal arrivals are an important factor in the perception of spatial images [14], due to psychoacoustic factors such as the precedence effect [15, 16] where late arriving signals are given less prominence than earlier ones when evaluating direction. When two identical stimuli are presented from different loudspeakers with a short time delay, precedence effect experiments have shown that as the time delay increases to 1 ms the image shifts to the direction of the leading sound. After this point the earlier arriving signal dominates and the image is localised from the direction of the leading sound. If the time delay is further increased then both signals will be heard as distinct sounds. The point at which this happens is known as the echo threshold.

The listening experiment in [13] was repeated for an array radius of 5 m in order to test the influence of increased time delays on the localisation of ambisonic images. The listeners were the same proportional distance from the centre as in the first experiment so the relative loudspeaker gains would be approximately equal in both experiments and the time difference between different loudspeaker signal arrivals would increase. In [13] the arrival time delays had a range up to 4 ms (beyond the 1 ms localisation dominance threshold) and yet it was still possible to localise the ambisonic image when listening at the off-centre positions. For the 5 m array the range increases to approximately 9 ms (given in table 1) so it might be expected that the increased time delays mean the later signals will contribute less to the localisation of the image, increasing the difficulty and impacting on the result. This paper presents the results for the large loudspeaker array and compares the results for both array sizes to investigate the influence of the increase of arrival time differences.

The paper is ordered as follows: section 2 presents the outline of the experiment conditions and the task, section 3 gives the results for the large (5m) loudspeaker array experiment, section 4 compares the results of the two array radii and section 5 provides a discussion relating to the comparison of results.

2. EXPERIMENT PROCEDURE

A detailed description of the experiment procedure can be found in an earlier paper [13] which presents the results of the experiment for a 2.2 m loudspeaker array.

In [13] the experiment took place in an 8-channel studio in the Sonic Arts Research Centre (SARC) (critical distance of 1.8 m) whereas for the large array it took place in the Sonic Laboratory in SARC [17]. The Sonic Lab has dimensions of 17 m × 13 m × 14 m. The reverberation time is 0.4 s at 1 kHz, giving a critical distance of 5 m, when in its most acoustically damped conditions and additional acoustic treatment was added at several first reflection points.

A standard max r_E decoder was used for the two Ambisonics setups: first with four loudspeakers (o1spk4) and third order with eight loudspeakers (o3spk8). The loudspeaker placement is shown in figure 1. No near-field compensation [3] was used. The experiment was implemented and run in real time using Max/MSP.

The array radius was 5 m and the listeners were placed such that the ratio of their distance from the centre of the array and

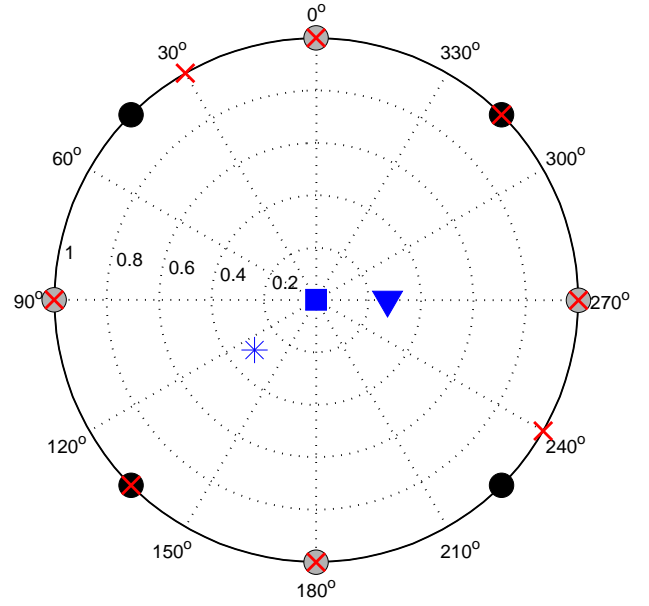


Figure 1: The 10 loudspeaker positions used in both experiments for the ambisonic arrays, the targets and the three listening positions. The eight loudspeakers used for the third order HOA array are circles and the four that were also used for the first order array are coloured black. Target positions are indicated by red crosses. The listening positions are shown as blue shapes around the centre. The radial distance has been normalised to show the relative distance the listeners were from the centre of the array in both cases.

the array radius was equal to that in the small array experiment. The off-centre positions were chosen to be approximately one seat away from the centre with the 2.2 m loudspeaker array and were scaled for the 5 m array. The layout of the target and array loudspeakers, along with the listening positions, is shown in figure 1. Both experiments took place in acoustically dampened conditions using 8 Genelec 1030a loudspeaker for the array, supplemented with 2 Genelec 1029a loudspeakers as targets. All of the loudspeakers were filtered to flatten their frequency responses and reduce the influence of the different spectral characteristics of the two loudspeaker models. The experiment with the 2.2 m array experiment and presented in [13] took place in February 2013 and the 5 m loudspeaker array took place in July 2013.

As the task, listeners were asked to move an ambisonic image pointer using a dial interface so that it appeared to be arriving from the same direction as a target sound produced by a single loudspeaker. The target and pointer sounds were played alternately to the listener. A maximum of 25 target-pointer repetition trials were allowed and listeners could end earlier if they were satisfied with their pointer placement. The target and system combinations were randomised and carried out for each of the three listening positions. The order of the listening positions was randomised for each listener to minimise the influence of fatigue.

A total of 12 subjects took part in the experiment, 1 woman and 11 men, with age ranging from 20 to 46. Five of the subjects self reported as experienced at listening to spatial audio rendering and 6 of them participated in the earlier experiment presented in [13]. None of the subjects reported any hearing loss but no formal

Table 1: The arrival time delays relative to the first arriving signal for each of the ambisonic array loudspeakers at the two off-centre listening positions.

	0°	45°	90°	135°	180°	225°	270°	315°
2.2 m array right	2.0 ms	3.1 ms	3.5 ms	3.1 ms	2.0 ms	0.7 ms	0.0 ms	0.7 ms
2.2 m array left-back	3.3 ms	2.0 ms	0.6 ms	0.0 ms	0.9 ms	2.4 ms	3.5 ms	3.9 ms
5 m array right	4.5 ms	7.0 ms	8.0 ms	7.0 ms	4.5 ms	1.5 ms	0.0 ms	1.5 ms
5 m array left-back	7.5 ms	4.6 ms	1.3 ms	0.0 ms	2.1 ms	5.5 ms	8.1 ms	8.8 ms

test was carried out before the experiment.

3. RESULTS OF LARGE ARRAY EXPERIMENT

3.1. Influence of Parameters

3.1.1. Number of Target-Pointer Trials Taken

Given that the number of trials of target-pointer pairs during each run was limited to 25 it is worth considering the proportion of target-pointer runs during which the user was still moving the pointer on the final trial. Over the whole experiment and all three listening positions, 6.42% of the runs had the pointer still moving on the final trial.

As was found in [13], the total number of trials needed (from 1 to 25) during each run follows the same trend across all three listening positions for both systems. The distribution of final pointer angles depends more on the order and system used rather than on the listening position.

Further investigation into each of these individual runs shows in which runs subjects were converging on a result and in which they were still making large movements when the scene ended. A run was deemed to be converging if the standard deviation of the pointer angle was less than 10° over the final 5 pointer trials and diverging if it was larger. It was found that a total of 63.1% of the runs that used all 25 trials were diverging. Furthermore, 26% for the centre listening position, 69% for the right and 51% for the left-back were diverging. The portion of the diverging runs that were using the o1spk4 system were 86%, 81% and 90% for the centre, right and left-back positions, highlighting again that the lower order system increased the difficulty.

3.1.2. Effect of Target Repetitions

Since each target was tested three times with each system and at each listening position the effect of these repetitions was tested to see if there was a learning effect. Using non-parametric analysis of variance (ANOVA) it was found that there was no statistical difference in the absolute error over each of the three repetitions ($\chi^2(2) = 0.51$, $p = 0.7751$). This suggests there was no learning or fatigue effect over the three repetitions.

3.2. Total Absolute Error

3.2.1. Influence of the System

Using a non-parametric ANOVA on the absolute error across all of the targets and listening positions, the o1spk4 and o3spk8 systems were found to be statistically different ($\chi^2(1) = 240.42$, $p < 0.01$). The median absolute error was larger for the first order system than the third order system (21.8° and 10.1° respectively), as was the interquartile range (27.4° and 13.4°), indicating the increased localisation performance for the higher order system.

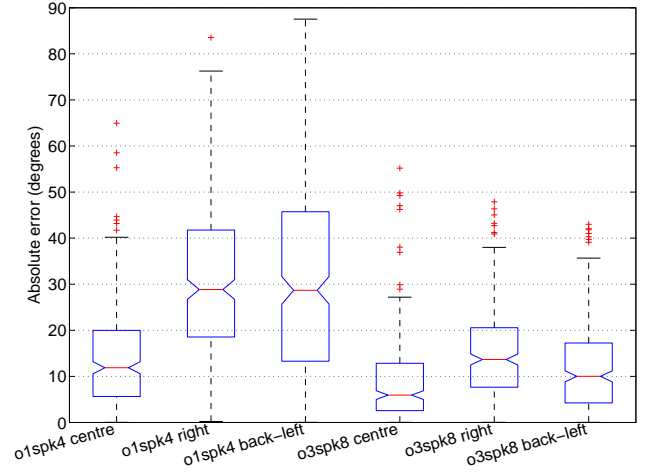


Figure 2: The median, interquartile and 95% values of the total absolute error for the two ambisonic systems at the three listening positions. The red crosses are the outlier values. For clarity some of these values were truncated for the o1spk4 right (1 point) and o1spk4 left-back (1 point).

3.2.2. Influence of the Listening Position

The absolute error at each of the different listening positions for all of the targets is higher for both the off-centre positions than it is for the centre, as shown in figure 2. This is true for both systems but the effect is much smaller with the o3spk8 system. A non-parametric ANOVA on the absolute error and listening position shows that there is statistical difference between the three listening positions ($\chi^2(2) = 173.3$, $p < 0.01$). Further analysis showed a significant difference between the two off-centre seats for the o3spk8 system ($\chi^2(1) = 15.09$, $p < 0.01$). However, the results for o1spk4 at the right and left-back positions are not statistically different ($\chi^2(1) = 0.05$, $p = 0.8155$).

Comparing the o1spk4 results for the centre and o3spk8 for the two off-centre seats in figure 2 shows that there is a comparable amount of overall error in all three. The median absolute error is 11.88°, 13.66° and 10.01° with an interquartile range of 14.35°, 12.91° and 12.99° for o1spk4 (centre), o3spk8 (right) and o3spk8 (left-back) respectively. There is no significant statistical difference between the o1spk4 centre or either of the off-centre o3spk8 systems so the global performance of these three system/listening position combinations can be said to be equivalent, although the two off-centre o3spk8 systems are significantly different to each other.

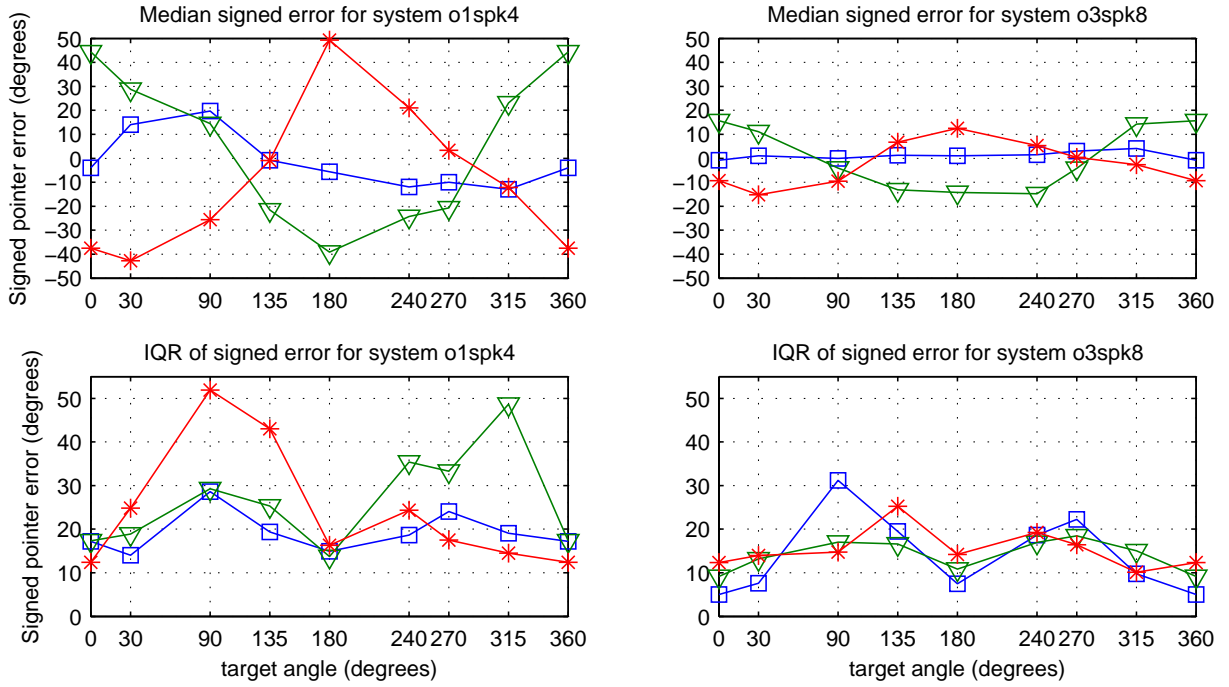


Figure 3: (Top) The median signed error for all three listening positions and the two ambisonic systems. Positive error indicates the pointer was shifted in an anti-clockwise direction from the target angle and negative error indicates a clockwise shift. (Bottom) The interquartile range for all three listening positions and the two ambisonic systems. Results for the centre are squares (blue in the coloured version), for the right are triangles (green in the coloured version) and for the left-back are stars (red in the coloured version)

3.3. Analysis of Pointer Error for Off-Centre Listeners

The signed error at each of the three listening positions indicates the magnitude and direction of the pointer results from the target angle. This is shown in figure 3 where a positive error indicates the pointer was shifted in an anti-clockwise direction from the target angle and negative error indicates a clockwise shift. For the off-centre listening positions the smallest shifts are for the targets that lie on or close to the line which passes through the listening position and the centre off the array. The pointer position is effectively shifted in a direction such that the pointer angle moves to the more distant side of the loudspeaker array.

The spread, indicated by the interquartile range, is also shown in figure 3. For the off-centre o1spk4 system results the spread is largest for the targets that are nearest to the listening position (such as those at 240°, 270° and 315° for the right position), suggesting the proximity to the loudspeakers meant a large range of pointer angles are likely to be perceived as coming from a smaller range of direction closer to the nearest loudspeaker. For the o1spk4 and o3spk8 systems at the centre listening position there is a very strong pattern where the highest spread is generally at the $\pm 90^\circ$ targets and lowest at 0° and 180°.

The spread of pointer angles for o3spk8 off-centre is actually comparable to or lower than o1spk4 at the centre for certain target positions (those nearest to the front and back). This could allow certain pointer positions to be more sharply defined for o3spk8 off-centre than o1spk4 can even manage at the centre.

4. COMPARISON OF SMALL AND LARGE ARRAY RESULTS

In this section a comparison is made between the results of the large 5 m loudspeaker array presented in the previous section and the small 2.2 m array presented in [13]. This is to investigate the influence of increasing differences between the loudspeaker arrival times and their influence on the localisation of ambisonic images.

4.1. Overall Absolute Error Between Experiments

A non-parametric ANOVA was carried out on the absolute error over all conditions between the results of both experiments. It was found that there was no statistically significant difference between the two experiments. This suggests that the overall performance for both array sizes was equal.

Testing for significance by splitting the results into the two orders show no significant difference between the two experiments for o1spk4 ($\chi^2(1) = 0.05$, $p=0.8212$) or o3spk8 ($\chi^2(1) = 0.02$, $p=0.8969$). The same is true when testing the overall error over both systems at each listening position between the two experiments: centre ($\chi^2(1) = 0$, $p=0.9625$), right ($\chi^2(1) = 2.14$, $p=0.1438$) and left-back ($\chi^2(1) = 1.66$, $p=0.1973$).

4.2. Listener Crossover

Of the 12 listeners who took part in the experiment presented here and the 15 that took part in [13] there were 6 that took part in both.

Table 2: The average intralistener variation for each system-seat combination for all listeners in both experiments.

Listener position	Small array		Large Array	
	o1spk4	o3spk8	o1spk4	o3spk8
centre	22.02°	17.27°	20.93°	16.22°
right	32.69°	20.70°	29.18°	18.28°
left-back	26.05°	16.85°	30.65°	16.00°

A non-parametric ANOVA was performed on the absolute error over all conditions for the results of the (a) small array experiment for those who didn't do the large array experiment (b) small array experiment for those that did both experiments (c) large array experiment for those who didn't do the small array experiment (d) large array experiment for those that did both experiments. It was found that there was no statistically significant difference between each of the groups ($\chi^2(3) = 6.13$, $p=0.1054$). This indicates that there was no difference in the overall performance between those who carried out both experiments (no learning effect from the earlier experiment) and those who did only one of the experiments.

4.3. Intralistener Variation

Given that there are no large differences between the two experiments in their overall results, it is interesting to look at the intralistener variation. The intralistener variation is defined here as the range of final pointer positions for the three repetitions of each condition for each listener. This will give some idea of the consistency with which the listeners were able to perform the task between the two loudspeaker arrays. When considering all listeners from both experiments it appears that for each system and listening position combination the intralistener variation between the six combinations was approximately equal between the two array sizes, as can be seen in table 2.

With the exception of o1spk4 at the left-back position, the large array has marginally lower average intralistener variation, though this might have been expected to be larger due to increased confusion from the greater arrival time delays. This suggests that, with the exception of the o1spk4 left-back position with the large array, there was no large decrease in the consistency with which each listener complete the task for the large array.

It is also worth noting that the lower overall error for the higher order system is mirrored in the intralistener variation as in each case the variation is lower for o3spk8 than o1spk4.

4.4. Global Signed Error

When comparing the results of the signed error between the small and large arrays, the overall trend, as shown in figure 3 for the large array, is the same for both the small array in [13] and for the large array. The mean absolute difference between the median results for the two experiments is shown in table 3. It shows that there is less error deviation for the o3spk8 system than o1spk4, indicating that at higher orders an ambisonic image is more consistent when the loudspeaker array size increases.

It is also evident that the mean absolute difference in the results for the centre listening position is of a similar size to or greater than the two off-centre listening positions. This perhaps suggests that the deviation had something to do with differences in the rooms the two experiments took place in rather than the increased time differences for the off-centre positions, otherwise one

Table 3: The mean absolute difference between the median results and interquartile ranges for the small and large array experiments.

Listener position	Difference in medians		Difference in IQRs	
	o1spk4	o3spk8	o1spk4	o3spk8
centre	7.24°	2.31°	3.61°	2.91°
right	7.03°	3.82°	4.31°	4.49°
left-back	5.55°	2.31°	4.96°	5.65°

Table 4: The median and interquartile range of the final pointer angles for the significantly different o3spk8 target and listening position combinations.

Target angle and listener position	Small array		Large array	
	median	IQR	median	IQR
30° centre	26.9°	9.4°	31.0°	7.6°
0° right	9.7°	9.3°	15.7°	9.2°
30° right	36.8°	6.8°	41.1°	13.2°

might have expected larger differences for the off-centre positions than the centre.

When comparing the interquartile range of both experiments there is a greater difference between the two trends. This suggests that while the overall median results stay relatively stable between both experiments the distribution is slightly altered. However, the mean absolute difference between the IQRs (table 3) for o1spk4 is similar for all three listening positions, with no larger variation seen for the off-centre seats. For o3spk8 the difference is slightly lower for the centre position by just more than a degree, indicating a greater similarity between the two experiments.

4.5. Target Positions

It has been shown that there is no significant difference between the two experiments when considering the absolute error at each seat or for each system the results for individual targets will be investigated to determine the similarity of the results in greater detail.

Carrying out a non-parametric ANOVA on the signed error of the two experiments for each target at each seat for the two systems shows that 12 out of 48 of the combinations are statistically different between the two experiments at a significance level of $p<0.05$.

Of the significantly different targets 9 are with the o1spk4 system and the results of the final pointer angle are shown in figure 4. The significantly different results for o3spk8 are omitted for visual clarity but their median and IQRs are presented in table 4.

The fact that the results of 75% of the all target, system and listening positions combinations are statistically similar shows a remarkable consistency between the two array sizes. 7 of the significantly different target, system and listening positions combinations (5 were o1spk4) were at the off-centre listening positions. That 5 out of 12 of the significantly different results occur for the central listening position suggests that the different rooms for the two experiments had an influence, since other than that the conditions were the same for both experiments. It is also possible that an increase in the number of participants might have caused the results to converge. It is therefore difficult to know if the significant differences at the off-centre positions are for similar reasons as those at the centre or if they represent a genuine significant difference caused by the differences in the arrival time delays.

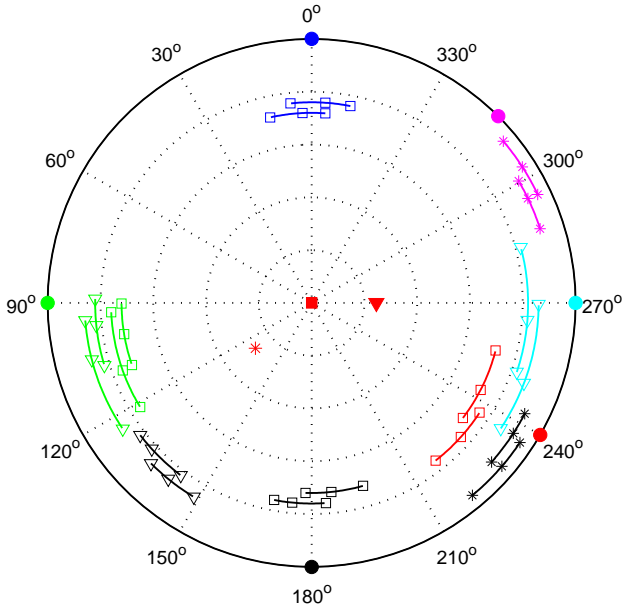


Figure 4: The 9 median and interquartile values pointer results for the o1spk4 system that are significantly different between the 5 m and 2.2 m loudspeaker array experiments. Results for the centre are represented by a square, results for the listening position represented by a triangle and those for the left-back are a star. The colour of the results matches the colour for each of the targets, shown as circles on the circumference of the plots. The inner of the pair of results corresponds to the 2.2 m experiment results and the outer to the 5 m array experiment.

4.6. Informal Comments on Image Timbre and Splitting

Similar to the comments from [13], with the o1spk4 system at off-centre listening positions, subjects reported a change of timbre. A change in timbre was not commented upon for the centre listening position. This is in line with Solvang [18] who showed that Ambisonics will be spectrally impaired outside the central listening position. The pointer timbre when the source was on the nearer side to an off-centre listener was different to that on the far side of the array. The change in timbre was also associated with change in pointer diffuseness when moved around the array.

In the small array experiment, two subjects reported occasionally being able to hear two signals when sitting in the off-centre position, one in the pointer direction and another from a nearby loudspeaker. In the large array experiment this was reported by many of the subjects for the o1spk4 system. It was reported to happen when the pointer was on the far side of the array (i.e. the direction of the 90° target when at the right listening position and the 315° target at the left-back position) rather than when it was nearby. Nettingsmeier and Dohrmann [19] tested an array of average radius 8.7 m and although no formal testing was carried out they found no echoes when on the far side of the array to the ambisonic image when using pink noise at third order. The time delays in this case were up to 36 ms, far beyond those in either experiment presented here.

5. DISCUSSION

The previous section showed that the results of the large and small array are remarkably similar between the two experiments when considering factors such as the signed and absolute error, and intralister variation. In both experiments the pointer angle was effectively shifted to the further side of the loudspeaker array. As discussed in [13], this effect is attributed to the proximity of the nearer loudspeakers, which have earlier arrival times and increased gains relative to the centre. Only 12 out of a possible 48 target and listening position combinations were statistically different and given that 5 of these were for the centre listening position suggests that the largest source of differences between the two experiments may be that they were performed in different rooms rather than due to the change in array size. Some of the variation at the centre seat might also have been caused by the fact that a max r_E decode is suboptimal for the centre position [20] and therefore could have lead to instability in the localisation results in this position. An increase in the number of participants in the experiment might have shown the difference to converge to zero.

The arrival time delays at off-centre positions were up to 9 ms for the large array experiment. The arrival time delays for each of the array loudspeakers relative to the first arriving signal is shown in table 1. Precedence effect [15, 16] experiments with two signals, a lead and a lag, suggest that the later arriving signal's contribution to image localisation will be suppressed if it arrives more than about 1 ms after the lead. However, for the 5 m ambisonic array all of the delayed signals arriving at the off-centre listening positions are beyond this 1 ms threshold but still influence the image direction. The Haas effect [21], where increasing the gain of the delayed signals significantly above that of the first arriving one can override the precedence effect. This is likely one factor that allows the image to move around the array and avoids it remaining fixed in the nearest loudspeaker. Another possible factor is that classic precedence effect experiments use two loudspeakers at $\pm 45^\circ$ but in the setup used in the experiments presented here the signals arrive from several directions, often without symmetry. Goupell et al. [22] also showed that the addition of a third reflection to a precedence effect test can cause more complex interaction between the signals than explained by the lead-lag model.

The similarity of the results between the two experiments suggests that there is a greater reliance on the loudspeaker gains than on the arrival time differences. This reflects the result by Peters [23] where a quantitative investigation found “that off-center sound degradation is primarily caused by the level differences of the loudspeaker feeds”, rather than the time differences, when testing a 5.1 system at off-centre positions. There may be some perceptual adaptability to changing time delays between the loudspeaker signal arrivals. Adaptability to an enclosed environment has been suggested by Hummersone et al. [24] in the context of the precedence effect in rooms of different sizes, suggesting that echo suppression parameters used by the spatial auditory system adapt to their environment. Something similar could be happening with the ambisonic reproduction, allowing the arrival time delays to increase without greatly adversely influencing the localisation.

The similarity in the arrival gains of at least two of the loudspeakers for o1spk4 when the image is placed on the far side of the array, coupled with the relatively large arrival time differences, means that the image has the potential to split if the time delays are greater than the echo threshold [16]. This splitting is not exhibited for the o3spk8 system because the later arriving signals are

of a much higher relative gain compared with the early signals. In his thesis, Frank [4] found that for an off-centre listener there was image splitting for a third order, basic decode image. This splitting was not exhibited for the same system with a max r_E decode. This was attributed to the more prominent side-lobes of the basic decode. A similar effect is most likely occurring here for the o1spk4 system.

6. CONCLUSION

A reverse acoustic target-pointer experiment was carried out to investigate the influence of off-centre listener positioning on localisation with first and third order ambisonics. The experiment was carried out on a 5 m radius loudspeaker array and the results are compared to earlier ones obtained from a 2.2 m radius array.

The first order system had higher overall error than the third order one, as shown in previous localisation experiments [13, 25, 7, 8]. The pointer was shifted toward the further side of the loudspeaker array to compensate for the proximity of the nearest loudspeakers, which have both earlier arrival times and increased gains relative to the centre position.

Remarkably similar results to those in [13] were found. The overall performance, measured by the absolute error, of both systems were the same for each system and listening position combination for both array radii. A comparison of the results for each of the individual targets with each system and listening position combination showed no statistical difference for 36 out of the 48 possible combinations. This suggests that the increase in arrival time delays due to the increase in the array size is a less significant factor than the loudspeaker gains. Of the 12 combinations which were statistically different, 5 of them were for the centre listening position. This suggests that some other factor, such as the rooms the experiments took place in, might explain the difference in the conditions for the off-centre seats, rather than the differing arrival time differences between the two array.

Acknowledgement

Many thanks to those who took the time to participate in the test and without whom this paper would not have been possible.

7. REFERENCES

- [1] M. A. Gerzon, "Periphony: With-Height Sound Reproduction," *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 2–10, January/February 1973.
- [2] J. Daniel, *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*, Ph.D. thesis, Université Pierre et Marie Curie (Paris VI), 2000.
- [3] J. Daniel, R. Nicol, and S. Moreau, "Further investigations of high order ambisonics and wavefield synthesis for holographic sound imaging," in *114th AES Convention, Amsterdam*, 2003, pp. 1–18.
- [4] M. Frank, *Phantom Sources using Multiple Loudspeakers in the Horizontal Plane*, Ph.D. thesis, University of Music and Performing Arts Graz, 2013.
- [5] M. A. Gerzon, "Design of Ambisonic Decoders for Multispeaker Surround Sound," in *58th AES Convention*, New York, NY, 1977.
- [6] D.G. Malham, "Experience With Large Area 3-D Ambisonic Sound Systems," in *Proceedings of the Institute of Acoustics Autumn Conference on Reproduced Sound*, 1992, vol. 14, pp. 209–216.
- [7] M. Frank, F. Zotter, and A. Sontacchi, "Localization Experiments Using Different 2D Ambisonics Decoders," in *25th Tonmeistertagung - VDT International Convention*, Leipzig, 2008.
- [8] E. Bates, G. Kearney, F. Boland, and D. Furlong, "Localization Accuracy of Advanced Spatialization Techniques in Small Concert Halls," in *153rd Meeting of the Acoustical Society of America*, Salt Lake City, 2007.
- [9] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, June 1997.
- [10] V. Pulkki, "Uniform Spreading of Amplitude Panned Virtual Sources," in *Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, 1999, pp. 187–190.
- [11] H. Wierstorf, Sa. Spors, and A. Raake, "Wahrnehmung künstlich erzeugter Schallfelder," in *German Annual Conference on Acoustics (DAGA)*, Oldenburg, Germany, 2014.
- [12] A. J. Berkhout, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764 – 2778, 1993.
- [13] P. Stitt, S. Bertet, and M. van Walstijn, "Perceptual Investigation Of Image Placement with Ambisonics For Non-Centred Listeners," in *16th Intl. conference on Digital Audio Effects (DAFx-13)*, Maynooth, Ireland, 2013, pp. 1–7.
- [14] G.S. Kendall, "Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music," *Organised Sound*, vol. 15, no. 03, pp. 228–238, December 2010.
- [15] H. Wallach, E.B. Newman, and M.R. Rosenzweig, "The Precedence Effect In Sound Localization," *The American Journal of Psychology*, vol. 62, no. 3, pp. 315–336, July 1949.
- [16] R.Y. Litovsky, H.S. Colburn, W.A. Yost, and S.J. Guzman, "The Precedence Effect.," *J. Acoust. Soc. Am.*, vol. 106, no. 4 Pt 1, pp. 1633–1654, October 1999.
- [17] SARC, "Sonic Lab," <http://www.sarc.qub.ac.uk/sites/sarc/AboutUs/TheSARCBuildingandFacilities/TheSonicLab/SonicLabSpecs/>, last accessed: 09-04-2014.
- [18] A. Solvang, "Spectral Impairment for Two-Dimensional Higher Order Ambisonics," *J. Audio Eng. Soc.*, vol. 56, no. 4, pp. 267–279, 2008.
- [19] J. Nettingsmeier and D. Dohrmann, "Preliminary Studies on Large-Scale Higher-Order Ambisonic Sound Reinforcement Systems," in *2011 Ambisonics Symposium*, Lexington, KY, 2011, pp. 1–6.
- [20] M. Gerzon, "General Metatheory of Auditory Localisation," in *Audio Engineering Society Convention 92*, Vienna, March 1992, vol. 82, pp. 273–286.
- [21] H. Haas, "The Influence of a Single Echo on the Audibility of Speech," *J. Audio Eng. Soc.*, vol. 20, no. 2, pp. 146–159, March 1972.

- [22] M.J. Goupell, Gongqiang Yu, and Ruth Y Litovsky, “The effect of an additional reflection in a precedence effect experiment.,” *The Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 2958–67, April 2012.
- [23] N. Peters, *Sweet [re]production : Developing sound spatialization tools for musical applications with emphasis on sweet spot and off-centre perception*, Ph.D. thesis, McGill University, 2010.
- [24] C. Hummersone, R. Mason, and T. Brookes, “Dynamic Precedence Effect Modeling for Source Separation in Reverberant Environments,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 18, no. 7, pp. 1867–1871, September 2010.
- [25] S. Bertet, J. Daniel, E. Parizet, and O. Warusfel, “Investigation on Localisation Accuracy for First and Higher Order Ambisonics Reproduced Sound Sources,” *Acta Acustica united with Acustica*, vol. 99, pp. 642–657, July 2013.